Report of the April 2003 Meeting of the Science Archive Working Group

The SAWG held their third meeting on April 22-23, 2003 at NASA HQ, with the following members present: Julian Borrill, Joel Bregman (Chair), Roger Brissenden, Damian Christian, Menas Kafatos, Carol Lonsdale, Bill Oegerle (Deputy Chair), Tom McGlynn, Sally Oey, Rick White, and Jonathan Borden, along with the NASA HQ personnel Paul Hertz, Jeff Hayes, Alan Smale, and Joe Bredekamp.

The ADEC, and Virtual Observatory Activities

At our October 2002 meeting, the SAWG requested a white paper from the ADEC in which they describe a plan for improving connectivity and services between the archive centers. The motivation for this was that the SAWG believed that NASA science would benefit considerably from this effort, and that the archive centers were in a excellent position to begin work in this direction. We also believed that it was important to begin this work while the NSF-sponsored ITR VO (Virtual Observatory) program was in progress, as we expect synergy between the NSF and NASA efforts, which would likely save NASA time and resources. The wording of our request was:

"the SAWG suggests that it is an appropriate time for the archival centers to increase their interoperability in order to meet strategic goals and to prepare for NASA participation in the anticipated VO. In particular, this development of VO-related activities should be considered along the lines of a NASA Project that will support the primary goals of the SEUS and OS roadmaps, in concert with the data that would be collected from the envisioned missions. Project Requirements should flow from these considerations, and there should be a well-defined set of data standards, goals, milestones, staffing levels, and budgets along a three-year timetable with a nominal start date in FY04. A "white paper" would be the result of this planning. This is envisioned as a modest NASA-only program of limited scope in which the staffing and budget models should be described for both an optimum and a minimal program."

The request from the SAWG was for a three-year, NASA-only proposal and in the six months since this request, the ADEC responded with a thoughtful and exciting proposal at relatively modest cost.

The proposal from the ADEC, entitled the *NASA Celestial Navigator System*, offers tremendous capabilities to NASA investigators. The current services offered by individual archive sites were developed in support of NASA missions with extensive input from scientists who use these data. These services are very widely used and have become invaluable for the scientific endeavor. However, there are barriers between the archive centers that prevent scientists from using these services across the archives. While the archive centers have taken modest steps to bring down these barriers, they do not have the existing resources to build the conduits between data sets housed at the various sites. Removing these barriers, consolidating databases, and providing a uniform interface lie at the center of the proposed *Celestial Navigator* System. The *Celestial Navigator* is not "The Virtual Observatory" that is all things to all people, but it is a significant and necessary step in building toward such a goal.

The SAWG endorses this proposal and believes that its execution is well within the current scope of technological abilities. Also, it leverages off the extensive experience with data, databases, and software at the archive sites as well as the NSF-ITR effort. The management and timelines seem sensible and there is enthusiasm by the member institutions, which should help the project to succeed. We believe that it is valuable to begin this program in FY04 as it will permit the project to be carried out concurrently with the NSF-ITR (which runs through FY06), with whom they would naturally interact (many of the members of the team for *Celestial Navigator* are participants in the NSF program). This NSF effort has developed a lot of "under the hood" elements, while the proposed NASA effort is a practical realization within the virtual observatory concept.

The costs beyond the end of this program are likely to be small. The software that is developed for Celestial Navigator will need to be maintained, but we estimate that it is at the 5-10% level of the yearly cost of development.

LAMBDA

The WMAP data were released on schedule and without problems, an impressive achievement since many members of the same team were responsible for the simultaneous release of the scientific results. These data are generally available in HEALPix format. The SAWG recognizes the emergence of the HEALPix pixelization as a de facto standard for *WMAP* and other CMB datasets. We are concerned that software enabling the analysis of data in this format should be made easily available to the community. As funds permit, *LAMBDA* should work with the developers of the HEALPix software to ensure easy community access to needed HEALPix libraries along with other software. *LAMBDA* might also explore the development of FITS WCS standards which would allow direct incorporation of HEALPix format data in FITS.

AISRP

The SAWG believes that the goal of the AISRP program is an important one, the support of software development that benefits a range of NASA scientists yet is beyond the software produced by archive or data centers. This program supports areas of special expertise (e.g., wavelet analysis, CMB analysis), adventurous prototypes that may evolve into valuable new tools, and other programs. Projects that do not produce deliverables were thought to be less valuable and this should be considered in the evaluation of future proposals.

Some components that will be included in a VO effort like *Celestial Navigator* were developed within the AISRP, so this might appear to be a natural source of funding for this effort. However, current AISRP funding is extremely modest and these funds are completely inadequate to support an effort such as *Celestial Navigator*, which will require new resources.

Other Business

There were a few other presentations and items that did not lead to "action items". Arnold Rots gave a presentation of CXC activities where the ongoing activities appear to be running smoothly. They have plans to produce Level 3 products, such as point source catalogs from all archival data, which should prove to be very useful, as comparable catalogs from other X-ray telescopes (e.g., *ROSAT*) are widely used. The status of *SOFIA* was presented by Tom

Greene, including instrument development, pipeline software production, and their PDMP. They seem to be proceeding sensibly, although some details about the archive and the interface have not yet been worked out.

A detailed plan for the NASA Herschel Science Center was submitted by George Helou, but there was insufficient time to review the document prior to the meeting. There will be an ongoing discussion between committee members through email and this issue will be taken up in detail during our next meeting.

There was a brief presentation on PDMP policy by Alan Smale in which the SAWG was asked if they should review PDMPs. Such plans should be reviewed and the SAWG seems to be a reasonable group to conduct such reviews as they arise. Alan Smale, who is new to code SZ, is known to the community through his many contributions in X-ray astronomy and his appointment improves the manpower situation within MO&DA, permitting him and Jeff Hayes to provide the attention needed to a broad range of issues.

Finally, we received feedback on previous suggestions. We had been concerned about the way in which GLAST planned to make proprietary the concepts behind winning proposals rather than data. The GLAST program has modified this and there will be no proprietary period. Also, we discussed with the GALEX representative about having an early release of calibration data and they have responded positively and are also modifying their GO program.

The next meeting of the SAWG is expected to take place in October 2003 and we welcome suggestions from the SEUS and the OS for future topics to be addressed.